

University of Groningen

## Controlling omitted variables and measurement errors by means of constrained autoregression and structural equation modeling

Suparman, Yusep

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*  
2015

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Suparman, Y. (2015). *Controlling omitted variables and measurement errors by means of constrained autoregression and structural equation modeling: Theory, simulations and application to measuring household preference for in-house piped water in Indonesia*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen.

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Samenvatting

Het eerste en hoofddoel van dit proefschrift is meer helderheid te brengen in de voorkeuren voor in pandig leidingwater in landelijke en stedelijke gebieden in Indonesië. Om dit doel te bereiken richt het proefschrift zich op de volgende objectieven. Ten eerste, omdat in het beschikbare gegevensbestand voor de analyse van dit hoofddoel – zoals trouwens ook in menig ander sociaal-wetenschappelijk gegevensbestand – een aantal systematische verklarende variabelen ontbreken, heb ik gerestricteerde autoregressie (CAR) en structurele vergelijkingen modellering (SEM) – ASEM – geïntroduceerd als een gecombineerde methodologie ter controle van zowel tijdvariërende ontbrekende variabelen als meetfouten. Ten tweede heb ik CAR vergeleken met de gebruikelijke alternatieve methoden voor de controle van tijdvariërende weggelaten variabelen, te weten: latente gefixeerde effectenmodel, eerste-orde verschillen regressie, regressie met gecentreerde variabelen en het autoregressieve model. Ten derde heb ik een likelihood-ratio toets voor CAR ontwikkeld bij aanwezigheid/afwezigheid van weggelaten variabelen. Tenslotte heb ik ASEM toegepast om de voorkeur voor in pandig leidingwater in landelijk en stedelijk Indonesië te analyseren.

**Doelstelling 1: Constrained autoregression – structural equation modeling (ASEM) to control time-varying omitted variables and measurement error.**

Met betrekking tot het eerste doel vond ik dat ASEM goede resultaten oplevert voor de controle van tijdvariërende weggelaten variabelen en dus een goede aanvulling vormt op standaard econometrische procedures zoals eerste-orde verschillen regressie. Een ander belangrijk voordeel is dat ASEM geen externe informatie behoeft om voor

meetfouten te controleren . SEM als methode om de meetfouten te controleren is goed gedocumenteerd in de literatuur (zie onder meer Jöreskog en Sörbom, 1996, en Bollen, 1989). De toepassing op stedelijk Indonesië in Hoofdstuk 2 laat zien dat de problemen van weggelaten variabelen en van meetfouten in verklarende variabelen gelijktijdig aangepakt moeten worden, zoals in ASEM.

**Doelstelling 2: The performance of CAR compared to the commonly used alternative methods to control time-varying omitted variables, viz. the latent fixed effect model, first order differencing regression, demeaning regression, and the autoregressive model.**

De analyse met betrekking tot deze doelstelling was Monte Carlo simulatie. De belangrijkste resultaten zijn als volgt. Ten eerste presteren de tijdvariërende benaderingen beter dan de tijdinvariante in termen van bias en gemiddelde kwadratische fout. Een tweede resultaat is dat autoregressie minder goed presteert dan gerestricteerde autoregressie met betrekking tot zo goed als alle simulatie parameter waarden. Op de derde plaats toonde analyse van de effecten van de simulatie parameters op de bias aan dat voor de tijdinvariante correctie-procedures en voor de autoregressieve benadering de bias die overblijft in de schatter van een regressiecoëfficiënt na controle voor de weggelaten variabele, een ingewikkelde functie is van met name de covarianties tussen de opgenomen variabelen en de weggelaten variabele en van de autoregressie-coëfficiënten van de weggelaten variabele en van de regressors. Alleen in het geval van CAR is de invloed van de determinanten van deze bias erg klein.

**Doelstelling 3: A likelihood ratio (LR) test of CAR in the presence/absence of omitted variables.**

Deze doelstelling werd ook geanalyseerd door middel van Monte Carlo simulatie. Het belangrijkste resultaat was dat de kans op een Type I fout van een CAR model

specificatie zonder weggelaten variabele (NOV) praktisch overal minder is dan het vooraf gespecificeerde significantie niveau van 0.01, terwijl het onderscheidingsvermogen van de CAR toets gebaseerd op de correcte hypothese van weggelaten variabelen overal 1 is. Een andere uitkomst is dat de toets zeer gevoelig is voor de autoregressie van de weggelaten variabele. De belangrijkste conclusie van Hoofdstuk 4 is dat de likelihood ratio toets van NOV versus CAR effectief kan worden toegepast om te beslissen over de toepassing van wel of geen CAR.

**Doelstelling 4: Analysis of the preference for in-house piped water in rural and urban Indonesia.**

De analyse van deze doelstelling is gebaseerd op de Indonesia Family Life Study (IFLS). Het belangrijkste resultaat was dat huishoudens in stedelijke en landelijke gebieden bij benadering dezelfde bereidheid om te betalen (WTP) laten zien voor in pandig leidingwater, gemeten als deel van de maandelijkse huishouduitgaven. Meer in het bijzonder, in pandige waterleiding verhoogt de gemiddelde huishuur met 34.24%, wat overeenkomt met respectievelijk 5.18% en 3.22% van de maandelijkse huishouduitgaven (voedsel en andere zaken) in stedelijke en landelijke gebieden. Het verschil tussen stedelijke en landelijke gebieden werd getoetst in een meervoudige steekproef analyse. De toets kon de nulhypothese van geen verschil tussen stedelijke en landelijke gebieden niet verwerpen. Voor de 25% huishoudens met de laagste uitgavenniveaus waren de huishuur schattingen in stedelijke en landelijke gebieden respectievelijk 9.41% en 7.57%. Deze uitkomsten zijn hoger dan die uit eerdere onderzoeken, waarin werd onderzocht wat arme huishoudens in Indonesië feitelijk uitgaven aan geven drinkwater. Mogelijke verklaringen voor de toename zijn: (i) leidingwater heeft een betere kwaliteit en beschikbaarheid dan geven drinkwater, (ii) de afgelopen 20 jaar is men zich meer bewust geworden van het nut van leidingwater, (iii)

de bevolking is beter geïnformeerd geraakt over de gevaren van het drinken van verontreinigd water.

### ***Methodologische aanbevelingen***

In dit proefschrift heb ik laten zien dat ASEM een effectieve methode is om te controleren voor weggelaten variabelen en meetfouten. In het bijzonder blijkt het CAR onderdeel van ASEM superieur te zijn ten opzichte van een verscheidenheid aan alternatieve procedures, die normaliter worden toegepast om de bias als gevolg van weggelaten variabelen te reduceren. Een ander voordeel is dat geen extra informatie is vereist. De beslissing om CAR te nemen tegenover een model zonder controle van weggelaten variabelen kan effectief worden getoetst door middel van een likelihood ratio toets. Een verder voordeel van CAR is dat het kan worden gecombineerd met SEM en zo toelaat om gelijktijdig zowel voor weggelaten variabelen als meetfouten te controleren. Het gebruik van ASEM en CAR in empirische onderzoek wordt sterk aanbevolen.

### ***Beleidsaanbevelingen***

De schattingen van de WTP-bedragen die in dit proefschrift zijn gevonden, impliceren dat stedelijke en landelijke huishoudens bij benadering dezelfde voorkeur hebben voor leidingwater. In het bijzonder zou de regering van Indonesië haar politiek moeten herzien om voorrang te geven aan stedelijke gebieden. De bevindingen in dit proefschrift tonen aan dat er zowel in de landelijke als stedelijke gebieden sterke voorkeur is voor in pandig leidingwater en dat beide gebieden in het beleid ouden moeten worden opgenomen.

De schattingen tonen ook aan dat de gemiddelde WTP in arme huishoudens behoorlijk boven het “betaalbaarheids criterium” van 5% van het totale inkomen ligt. Bovendien is vergelijkswijs de WTP van arme huishoudens hoger dan die van de rijkere huishoudens. In overeenstemming hiermee zouden zij prioriteit moeten krijgen. Mocht de regering besluiten tot een volledig dekkingsschema, dan zouden hoge watertarieven resulteren die de arme huishoudens weer kunnen afschrikken om aangesloten te worden (Jalan et al., 2009). In plaats daarvan zouden zij uitwijken naar minder veilige maar betaalbare voorzieningen. De regering zou daarom haar beleid moeten heroverwegen, in het bijzonder ten aanzien van subsidies, waar die al zijn, en deze invoeren, waar ze nog niet zijn (zie ook van den Berg en Nauges, 2012). Als alternatief zou de regering kunnen overwegen een geblokt tarief systeem op te zetten met het eerste blok gratis voor de armsten in de samenleving (Majumdar et al., 2009).



# **Stelingen Behorende bij het Proefschrift**

## **Controlling Omitted Variables and Measurement Errors by Means of Constrained Autoregression and Structural Equation Modeling Theory, Simulations and Application to Measuring Household Preferences for In- House Piped Water in Indonesia**

**Yusep Suparman**

1. Methodology, like sex, is better demonstrated than discussed, though often better anticipated than experienced. (Leamer, 1983)
2. The primacy of empirical modeling is the theoretical model which drives data collection and data analysis subject to methodological and statistical constraints.
3. Omitted variables and measurement errors can be adequately and simultaneously controlled by means of a constrained autoregressive - structural equation model (ASEM). (This thesis)
4. The constrained autoregression model outperforms the time-invariant omitted variable approaches and the autoregression model in terms of bias and mean square error. (This thesis)
5. Based on the probability of a Type I error and the power of a test, the likelihood ratio test is an adequate test to decide about the adoption of constrained autoregression to control omitted variables. (This thesis)
6. In-house piped water is the most effective mode for controlling waterborne diseases. (Hutton and Haller, 2004)
7. The preference for in-house piped water is inversely related to the availability and the quality of drinking-water from rivers, lakes and wells. (This thesis)
8. In contrast to the Government of Indonesia's water provision policy, households in rural and urban Indonesia virtually have the same preference for in-house piped water, as measured by their willingness to pay relative in terms of imputed house rent. (This thesis)
9. Doing a PhD is not only training one's expertise and skills but also shaping one's attitudes to be more open and more critical, more patient and more persistent.



10. *Cikaracak ninggang batu, laun-laun jadi legok* (Slowly but surely, small continuous dropping water will make a cavity in a stone): Patience and persistence are two important keys to success. (Sundaness proverb)